

# Entanglement Entropy in Quantum Gravity and the Plateau Problem

Dmitri V. Fursaev

*Dubna International University  
and  
the University Centre,  
Joint Institute for Nuclear Research  
141 980, Dubna, Moscow Region, Russia*

E-mail: fursaev@theor.jinr.ru

## Abstract

In a quantum gravity theory the entropy of entanglement  $S$  between the fundamental degrees of freedom spatially divided by a surface is discussed. The classical gravity is considered as an emergent phenomenon and arguments are presented that: 1)  $S$  is a macroscopical quantity which can be determined without knowing a real microscopical content of the fundamental theory; 2)  $S$  is given by the Bekenstein-Hawking formula in terms of the area of a co-dimension 2 hypersurface  $\mathcal{B}$ ; 3) in static space-times  $\mathcal{B}$  can be defined as a minimal hypersurface of a least volume separating the system in a constant time slice. It is shown that properties of  $S$  are in agreement with basic properties of the von Neumann entropy. Explicit variational formulae for  $S$  in different physical examples are considered.

# 1 Introduction

The concept of gravity as an emergent phenomenon [1] continues to be one of the attracting and stimulating ideas in theoretical physics [2], [3]. According to this point of view, the space-time metric is an effective low-energy variable and metric changes are collective excitations of underlying degrees of freedom of some unknown fundamental theory. The idea gets an inspiration from the condensed matter physics. For example, at very small scales the space-time may have a discrete lattice structure and gravitons, like phonons, may be collective excitations of the lattice. Gravity analogs in condensed matter physics may be very useful to develop an intuition about the physics at Planckian energies and to model phenomena predicted in quantum gravity [2].

The experience with condensed matter models tells us that macroscopical phenomena can be described without knowledge of the microscopical structure. Under certain conditions one can quantize the effective low-energy variables and get reliable results which are equivalent to computations in terms of genuine microscopical degrees of freedom [4]. This feature is also present in effective field theories, like the theory of pions (see [5] and references therein).

One of characteristics of a many-body system is a quantum entanglement which measures a degree of the correlation between different parts of the system. Quantum entanglement in condensed matter received a considerable attention in last few years. Its study helps to understand better collective effects in strongly correlated systems, see [6] for a review. Another application of the entanglement is a quantum information theory where it is used to store and process the information [7].

The quantum entanglement is quantified by an entropy. The present paper considers the entanglement entropy in a fundamental (quantum gravity) theory by adopting the point of view that gravity is an emergent phenomenon [1].

Let us start with the problem formulated in [8]: if the fundamental theory were known what would be the entropy of entanglement in a flat space-time between the degrees of freedom divided by a plane? According to [8] in the leading order approximation the entropy  $S$  for a sufficiently large but finite piece  $\mathcal{B}$  of the plane can be expressed as

$$S = \frac{\mathcal{A}(\mathcal{B})}{4G} \quad , \quad (1.1)$$

where  $\mathcal{A}$  is the area of  $\mathcal{B}$ . It follows from (1.1) that the main information about correlations of the degrees of freedom across the plane is encoded in the gravity coupling  $G$  in the low-energy sector of the theory. The conjecture was motivated by the black hole physics where the Bekenstein-Hawking entropy  $S^{BH}$  is given by (1.1) and  $\mathcal{A}$  is the area of the black hole horizon. One of possible interpretations [9]–[11] relates  $S^{BH}$  to an entanglement of the degrees of freedom across the horizon.

The aim of the present paper is to find a definition of the entanglement entropy  $S$  in the presence of the gravitational field and study its properties. The case considered in [8] is a system divided by a plane. In a flat space the plane is a minimal surface. We present a number of arguments that the entanglement entropy in a quantum gravity is given by (1.1) and in a static space-time the separating surface  $\mathcal{B}$  is a minimal least area hypersurface in a constant-time slice.

Finding a minimal surface for given boundary conditions is called the Plateau problem. The fact that  $S$  is a dynamical variable which has a simple geometrical meaning is rather

non-trivial from the physical point of view. It means that *the entanglement entropy in a quantum gravity theory can be measured solely in terms of macroscopical (low-energy) parameters without the knowledge of a microscopical content of the theory*. In this regard  $S$ , being a more general concept, is analogous to a thermodynamical entropy.

The relation between the entanglement entropy and minimal surfaces is not unexpected. In a different context, Ryu and Takayanagi [12], [13] suggested a holographic formula for the entanglement entropy in conformal field theories (CFT) which admit a dual description in terms of the anti-de Sitter (AdS) gravity one dimension higher. The holographic formula looks as (1.1) where  $G$  is a higher-dimensional gravity coupling. It yields the entropy of a CFT in terms of the area of a minimal surface embedded in a higher-dimensional bulk space-time. The holographic formula can be tested in case of two-dimensional CFT's and it reproduces correctly results of quantum computations.

Although our hypothesis and the suggestion of [12], [13] may be related they are different statements. The holographic formula is a geometrical representation of the entanglement entropy for a certain class of quantum field models. Thus, the entropy here is a model and cutoff dependent quantity.

The paper organized as follows. The arguments in favour of (1.1) are presented in section 2. We remind a path integral approach for computation of the entanglement entropy in QFT's and introduce a "partition function"  $\mathcal{Z}$  in a form of a Gibbons-Hawking path integral with a special choice of the boundary conditions. The entropy  $S$  can be computed from  $\mathcal{Z}$  in a statistical-mechanical manner. A similar construction was used in [14] to justify the holographic formula of [12], [13]. Quantum fluctuations of the geometry do not allow to fix the separating surface at the beginning. The surface is specified by boundary conditions. We show that the leading order approximation to  $\mathcal{Z}$  is determined by the geometries where the separating surface is a minimal co-dimension 2 hypersurface with a least volume.

The main suggestion about entanglement entropy in a quantum gravity theory is formulated in section 3 for static space-times. We first consider space-times where constant-time slices are Cauchy hypersurfaces with simply connected boundaries. Static space-times are singled out because the separating surface  $\mathcal{B}$ , which is an extremal co-dimension 2 hypersurface in the space-time, is a minimal co-dimension 1 hypersurface in a constant time slice. The latter feature ensures that  $S$  has the properties of the von Neumann entropy.

The Cauchy surface of a maximally extended Schwarzschild black hole space-time has two disconnected asymptotically flat regions. In section 4 we use this example to formulate the suggestion about the entanglement entropy in the presence of a black hole horizon. The bifurcation surface of the horizon is a minimal surface on a constant-time slice. One can relate to this surface an entanglement entropy which coincides with the Bekenstein-Hawking entropy and corresponds to the loss of information about states located behind the horizon.

The arguments of section 2 based on the path integral approach serve as a motivation. They cannot be considered as a real proof of (1.1) because they concern an effective theory. In section 5, we show that our suggestions regarding the entanglement entropy have an important feature: they are consistent with the subadditivity property of the von Neumann entropy. The material of this section is similar to the analysis of the entropy in CFT's with gravity duals [15], [16].

We give examples how to express changes of  $S$  solely in terms of macroscopical pa-

rameters of the system in section 6. The examples include changes caused by shifts of a point-like particle and by deformations of a cosmic string located near a separating surface. In section 7 we discuss quantum corrections to formula (1.1) from low-energy fields and suggest that ultraviolet divergences of an entanglement entropy encountered in QFT models are removed in the course of the standard renormalization of gravity couplings. The paper is ended with a discussion of its results in section 8.

## 2 Formulation of the problem

### 2.1 Entanglement entropy in QFT

We first recall a path-integral approach to calculation of entanglement entropy in QFT models [8]. Consider a stationary space-time  $\mathcal{M}$  with the number of dimensions  $d$  (the concrete value of  $d$  will not be important for our analysis). Let  $\partial_t$  be a time-like Killing vector field of  $\mathcal{M}$ . Suppose that the system is defined on a domain  $\Sigma$  of a constant time slice  $t = \text{const}$ .

We also assume that the system is in a thermal state with the temperature  $T$ . The corresponding density matrix is  $\hat{\rho} = \mathcal{N}^{-1} \exp(-\hat{H}/T)$  where  $\hat{H}$  is the Hamiltonian which is defined as a generator of the evolution of the given system along the Killing time  $t$ .

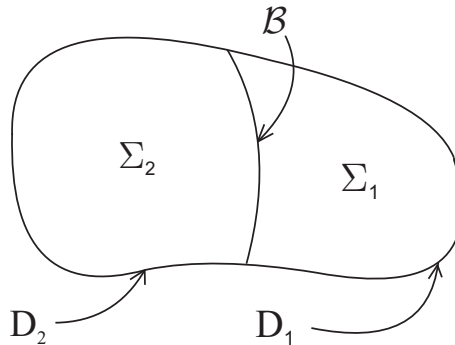


Figure 1: A constant time slice of a system spatially divided by a hypersurface  $\mathcal{B}$ .

Suppose that  $\Sigma$  is divided into two parts,  $\Sigma_1$  and  $\Sigma_2$ , by a co-dimension 2 hypersurface  $\mathcal{B}$  (see Fig. 1) and define a reduced density matrix for the region  $\Sigma_1$ ,

$$\hat{\rho}_1 = \text{Tr}_2 \hat{\rho} \quad , \quad (2.1)$$

where the trace is taken over the states located in the region  $\Sigma_2$ . The entanglement entropy is defined as

$$S_1 = -\text{Tr}_1 \hat{\rho}_1 \ln \hat{\rho}_1 \quad . \quad (2.2)$$

Analogously, one can define the entanglement entropy  $S_2$  in terms of the reduced density matrix obtained by tracing the degrees of freedom in  $\Sigma_1$ . There is a method of computation of  $S_1$  based on the relation

$$S_1 = -\lim_{n \rightarrow 1} \frac{\partial}{\partial n} \text{Tr}_1 \hat{\rho}_1^n \quad . \quad (2.3)$$

Consider the partition function of the given system

$$Z(T) = \text{Tr} e^{-\hat{H}/T} \quad .$$

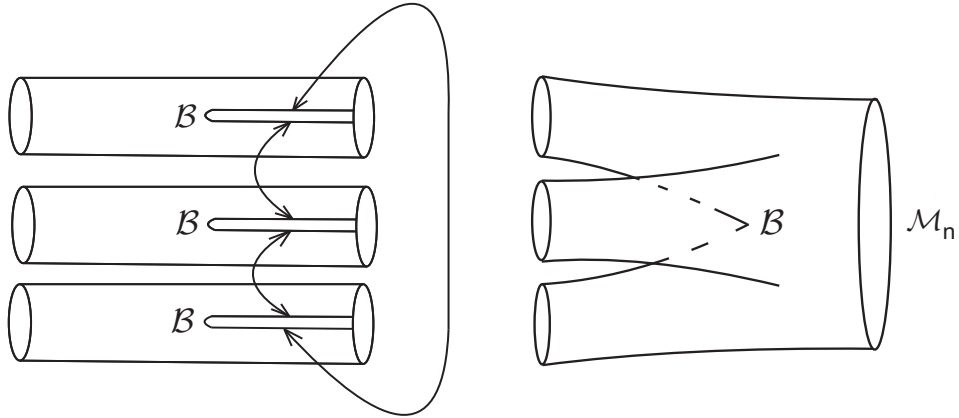


Figure 2: Construction of the space  $\mathcal{M}_n$  for a 2D finite-temperature system set on an interval  $\Sigma$ . For  $n = 3$  the space  $\mathcal{M}_n$  is obtained from 3 copies of the cylinders  $\Sigma \times S^1$ . The cylinders are cut and glued along the part  $\Sigma_1$  of a "constant time" slice.

According to the standard approach it can be written in terms of the path integral

$$Z(T) = \int [D\phi] e^{-I[\phi, T]} \quad , \quad (2.4)$$

where  $I[\phi, T]$  is the classical action for the given model. The fields  $\phi$  in (2.4) are defined on a Riemannian manifold  $\mathcal{M} = \Sigma \times S^1$  where  $S^1$  is a circle of the length  $T^{-1}$ .

The recipe for computing the entropy  $S_1$  is the following [8]. Consider a family of manifolds  $\mathcal{M}_n$  where  $n$  is a natural number and  $\mathcal{M}_1 \equiv \mathcal{M}$ . To obtain  $\mathcal{M}_n$  for  $n > 1$  one has to cut  $\mathcal{M}$  along one of its constant time hypersurfaces isomorphic to  $\Sigma$ . The cut is made along the domain  $\Sigma_1$  of  $\Sigma$ . Then one takes  $n$  identical copies of the manifolds and glues them along the cuts. This procedure is shown on Fig. 2 for a 2D theory. As a next step, for each  $n$  one defines a "partition function" on  $\mathcal{M}_n$

$$Z_1(\beta, T) = \int [D\phi] e^{-I[\phi, \beta, T]} \quad , \quad (2.5)$$

as a direct generalization of (2.4). The parameter  $\beta$  in (2.5) is  $2\pi n$ . The entropy can be represented in a statistical-mechanical form

$$S_1(T) = - \lim_{\beta \rightarrow 2\pi} \left( \beta \frac{\partial}{\partial \beta} - 1 \right) \ln Z_1(\beta, T) \quad . \quad (2.6)$$

The operation with  $\beta$  in (2.6) should be understood in the following way: one first computes  $Z_1(\beta, T)$  for  $\beta = 2\pi n$ , and then replaces  $\beta$  with a continuous parameter. This can be done even if  $\mathcal{M}_n$  itself cannot be defined at arbitrary  $\beta$ . Equation (2.6) coincides with the formula for the entropy in statistical mechanics if  $\beta^{-1}$  is interpreted as a temperature. Note that  $\beta^{-1}$  is a geometrical parameter, the physical temperature of the system is  $T$ . By cutting  $n$  copies of  $\mathcal{M}$  and gluing along  $\Sigma_2$  one can obtain the entanglement entropy in the region  $\Sigma_2$ .

The reduced density matrix (2.1) is given by a functional integral like (2.4) on  $\mathcal{M}$  with a cut along  $\Sigma_1$ . The values of the field variables on the both sides of the cut determine the "indexes" of the matrix  $\hat{\rho}_1$ . Thus, the trace like  $\text{Tr}_1 \hat{\rho}_1^n$  is determined by the path integral (2.5) and (2.6) follows from (2.3).

The method described above enables one to compute the entanglement entropy in quantum field models. The calculation can be done in the regularized theory in the presence of an ultraviolet cutoff. The dependence on the cutoff  $\varrho$  has a very simple form [9], [10]. For instance, the leading divergence in the entropy in a four-dimensional theory is proportional to  $A/\varrho^2$ , where  $A$  is the area of the separating surface  $\mathcal{B}$  ( $\varrho$  is assumed to have the dimension of a length). In condensed matter systems the entanglement entropy is finite and  $\varrho$  is a physical parameter associated, for example, with a lattice spacing.

## 2.2 Basic assumptions

To define the entanglement entropy in a quantum gravity theory one has to answer a number of questions. Here are some of them.

- 1) The notion of classical geometry may not hold at microscopical scales. Does the definition of a "separating surface" make sense in a quantum gravity theory?
- 2) Related to the first question is the problem of the entanglement of gravitational degrees of freedom. How can one take into account graviton contributions to the entropy?
- 3) The entanglement entropy of quantum fields is divergent. Can this problem be solved by a traditional prescription based on the renormalization? What are the physical constants which should be renormalized?

To approach these problems we assume that the gravity along with other low-energy fields (the fields of the Standard Model, for example) is an emergent phenomenon, in a sense that the fields represent an effective description of some other underlying fundamental variables. One can guess that it is the effective nature of the low-energy theory which results in ultraviolet divergences, while the underlying theory does not have this problem. We assume that the trace over the fundamental degrees of freedom which has to be done to compute the reduced density matrix in the quantum gravity is equivalent to the trace over low-energy variables including fluctuations of the metric.

In the present paper we restrict our analysis by theories at finite-temperatures, including vacuum states as a limiting case. This yields an opportunity to use Euclidean formulation of the theory. By following Gibbons and Hawking we work with a partition function for a canonical ensemble in a gravity theory [17], [18]

$$Z[h, \varphi] = \int [Dg][D\phi] \exp(-I[g, \phi]) \quad , \quad (2.7)$$

where  $I[g, \phi]$  is a classical Euclidean action for the metric  $g_{\mu\nu}$  and some set of matter fields  $\phi$ . The path integral (2.7) is considered as a low-energy approximation for a genuine partition function of a quantum gravity theory.

The "trajectories" in (2.7) are Riemannian manifolds  $\mathcal{M}$  and all possible field configurations living on them. The number of dimensions of  $\mathcal{M}$  is  $d$ ,  $d > 2$ . Two dimensional gravity models can be included in our analysis but they require a separate discussion.

The form of "trajectories" is specified by the boundary conditions. It is assumed that manifolds have a boundary  $\partial\mathcal{M} = \mathcal{T} = \mathcal{S} \times S^1$  where  $\mathcal{S}$  is a  $(d-2)$ -dimensional spatial boundary of the system. We demand that  $\mathcal{T}$  possesses a Killing vector field  $\xi = \partial_\tau$  with closed orbits  $S^1$ . The length of the circle  $S^1$  is fixed by the inverse temperature. It will be assumed that  $\mathcal{S}$  is a compact closed manifold. A typical example of the boundary is  $\mathcal{S} = S^2$ . The partition function depends on the metric  $h_{\mu\nu}$  of  $\mathcal{T}$  and on values  $\varphi$  of matter fields on  $\mathcal{T}$ .

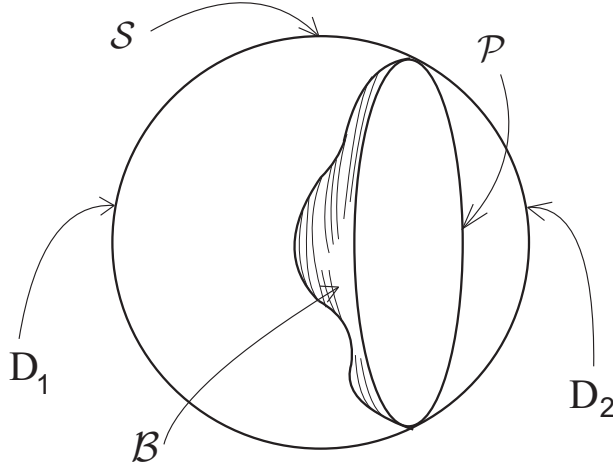


Figure 3: In a quantum gravity one can specify the entangled regions of a system by dividing its spatial boundary  $\mathcal{S}$  onto domains  $D_1$  and  $D_2$ . The separating surface  $\mathcal{B}$  is determined by a dynamical problem with the condition that  $\mathcal{B}$  ends at the boundary  $\mathcal{P}$  between  $D_1$  and  $D_2$ .

The action in (2.7) is

$$I[g, \phi] = I[g] + I_{\text{m}}[g, \phi] \quad , \quad (2.8)$$

$$I[g] = -\frac{1}{16\pi G} \int_{\mathcal{M}} d^d x \sqrt{g} (R - 2\Lambda) - \frac{1}{8\pi G} \int_{\partial\mathcal{M}} d^{d-1} x \sqrt{h} K \quad . \quad (2.9)$$

Here  $I_{\text{m}}[g, \phi]$  is an action of matter fields on  $\mathcal{M}$  whose explicit form is not relevant for further discussion. The boundary term in (2.9) is the Gibbons-Hawking term which depends on the trace of the extrinsic curvature  $K$  of  $\mathcal{T}$  embedded in  $\mathcal{M}$ . This term is needed to formulate a well-defined variational procedure.

In the semiclassical approximation the partition function takes the form

$$Z[h, \varphi] \sim \exp(-I[\bar{g}(h), \bar{\phi}(\varphi)]) \quad , \quad (2.10)$$

where  $\bar{g}(h)$  and  $\bar{\phi}(\varphi)$  are solutions to classical equations which define extrema of  $I[g, \phi]$  for given boundary conditions. Equations for the metric are the (Euclidean) Einstein equations

$$R_{\mu\nu} - \frac{1}{2} \bar{g}_{\mu\nu} R + \Lambda \bar{g}_{\mu\nu} = 8\pi G T_{\mu\nu} \quad , \quad (2.11)$$

with a matter stress-energy tensor  $T_{\mu\nu}$ . We denote by  $\bar{\mathcal{M}}$  a manifold whose metric is a solution to (2.11). It is assumed that the boundary condition  $\partial\bar{\mathcal{M}} = \mathcal{T}$  admits stationary solutions.

Because in a quantum gravity theory the metric except its boundary value cannot be fixed a priori, for the same reason the definition of entangled states of a system can be described only in terms of a partition of the boundary. We take a constant time ( $\tau = \text{const}$ ) slice  $\mathcal{S}$  of  $\mathcal{T}$  and consider division of  $\mathcal{S}$  by a hypersurface  $\mathcal{P}$ , see Fig. 3. The parts  $D_1$  and  $D_2$  of  $\mathcal{S}$  separated by  $\mathcal{P}$  will be treated as boundaries of two entangled regions.

Let  $S(D_1, T)$  be an entanglement entropy which appears when the degrees of freedom inside the region with the boundary  $D_2$  are traced out. By bearing in mind the method

described in the previous section we relate the entropy to a partition function

$$S(D_1, T) = - \lim_{\beta \rightarrow 2\pi} \left( \beta \frac{\partial}{\partial \beta} - 1 \right) \ln \mathcal{Z}(\beta, D_1, T) \quad . \quad (2.12)$$

The function  $\mathcal{Z}(\beta, D_1, T)$  is determined by a parameter  $\beta = 2\pi n$  where  $n$  is a natural number. We assume that  $\mathcal{Z}(\beta, D_1, T)$  can be represented in terms of a sum over geometries,  $\mathcal{M}_n$ , specified by a boundary condition  $\partial \mathcal{M}_n = \mathcal{T}_n$ .

The boundary space  $\mathcal{T}_n$  is constructed from  $\mathcal{T}$  as follows. One cuts  $\mathcal{T}$  along the part  $D_2$  of the slice  $\tau = \text{const}$  and glues  $n$  identical copies of  $\mathcal{T}$  along the cuts. Spaces  $\mathcal{T}_n$  have conical singularities on a surface where the cuts meet. The singular surface is isometric to  $\mathcal{P}$ . The structure of  $\mathcal{T}_n$  near the singular surface is  $\mathcal{C}_\beta \times \mathcal{P}$  where  $\mathcal{C}_\beta$  is a 2-dimensional cone with the circumference length  $\beta = 2\pi n$  around the tip.

For  $n > 1$  ( $\beta > 2\pi$ ) we define the partition function as:

$$\mathcal{Z}(\beta, D_1, T) = \sum_{\mathcal{B}} Z(\beta, D_1, \mathcal{B}, T) \quad , \quad (2.13)$$

$$Z(\beta, D_1, \mathcal{B}, T) = \int' [Dg][D\phi] \exp(-I[g, \phi]) \quad . \quad (2.14)$$

The path integral  $\int'$  in the right hand side of (2.14) is similar to (2.7) but it is taken over a class of special geometries,  $\{\mathcal{M}_n\}_{\mathcal{B}}$  (and field configurations living on them). The boundary of spaces from  $\{\mathcal{M}_n\}_{\mathcal{B}}$  is  $\mathcal{T}_n$ . Because  $\mathcal{T}_n$  have conical singularities  $\{\mathcal{M}_n\}_{\mathcal{B}}$  are also supposed to have conical singularities on a co-dimension 2 hypersurface  $\mathcal{B}$ . The surface  $\mathcal{B}$  has the following properties: it is located entirely in  $\mathcal{M}_n$  and intersects the boundary  $\mathcal{T}_n$  at  $\mathcal{P}$ . Near  $\mathcal{B}$  the spaces look as  $\mathcal{C}_\beta \times \mathcal{B}$ . Inside the class  $\{\mathcal{M}_n\}_{\mathcal{B}}$  the metric of  $\mathcal{B}$  is fixed.

To summarize, calculation of  $\mathcal{Z}(\beta, D_1, T)$  consists of two steps. One takes the path integral (2.14) over the class of spaces specified by the presence of the singular surface  $\mathcal{B}$  and then one sums contributions of different classes  $\{\mathcal{M}_n\}_{\mathcal{B}}$ . As a result, all possible geometries which obey the boundary conditions and have conical singularities on co-dimension 2 hypersurfaces are taken into account. The definition of  $\mathcal{Z}(\beta, D_1, T)$  is motivated by the form of the QFT partition function (2.5).

The summation symbol  $\sum_{\mathcal{B}}$  in (2.13) does not have a precise meaning, the situation common for path integrals. This is not a principle difficulty because we are dealing with an effective theory and our main result will be based on a semiclassical approximation.

## 2.3 Semiclassical approximation

To compute the integral (2.14) in the semiclassical approximation one has to take into account the form of the classical gravity action on  $\{\mathcal{M}_n\}_{\mathcal{B}}$ ,

$$I[g, \beta] = -\frac{1}{16\pi G} \int_{\mathcal{M}_n} d^d x \sqrt{g} (R - 2\Lambda) - \frac{1}{8\pi G} \int_{\partial \mathcal{M}_n} d^{d-1} x \sqrt{h} K - \frac{1}{8\pi G} (2\pi - \beta) \mathcal{A}(\mathcal{B}) \quad . \quad (2.15)$$

Here  $\mathcal{A}(\mathcal{B})$  is the volume of  $\mathcal{B}$ . The first term in the r.h.s. of (2.15) is an integral over a regular domain of  $\mathcal{M}_n$ , the last term appears as a result of conical singularities on  $\mathcal{B}$  [19]<sup>1</sup>.

---

<sup>1</sup>Conical singularities of the boundary space  $\mathcal{T}_n$  do not contribute to (2.15). This can be shown by using arguments presented in [14].



The semiclassical approximation to (2.14) is determined by extrema of the action (2.15) on  $\{\mathcal{M}_n\}_{\mathcal{B}}$ . Because the area of  $\mathcal{B}$  is fixed metric variations should vanish at  $\mathcal{B}$ . Thus, extrema of (2.15) are geometries which are solutions to gravity equations (2.11) outside  $\mathcal{B}$ .

Let  $\bar{\mathcal{M}}_n$  be such a solution for a class  $\{\mathcal{M}_n\}_{\mathcal{B}}$ . Any solution  $\bar{\mathcal{M}}_n$  with  $n > 1$  can be constructed from a solution  $\bar{\mathcal{M}}$  with  $n = 1$ . Consider to this aim an embedding of  $\mathcal{B}$  in  $\bar{\mathcal{M}}$ , such that  $\mathcal{B}$  ends on a constant time slice of the boundary  $\mathcal{T}$ . One cuts  $\bar{\mathcal{M}}$  along a co-dimension 1 hypersurface which starts at the domain  $D_2$  of the slice on  $\mathcal{T}$  and ends on  $\mathcal{B}$  (the precise choice of the surface does not matter). Spaces  $\bar{\mathcal{M}}_n$  are obtained by gluing  $n$  identical copies with the same cut so that to get conical singularities on  $\mathcal{B}$ . Because the cut intersects the slice of  $\mathcal{T}$  along  $D_2$  the constructed spaces also satisfy the required boundary conditions,  $\partial\bar{\mathcal{M}}_n = \mathcal{T}_n$ .

In the semiclassical approximation

$$-\ln Z(\beta, D_1, \mathcal{B}, T) \simeq I[\bar{g}(h), \beta] + I_m[\bar{g}(h), \bar{\phi}(\varphi), \beta] = nI[\bar{g}(h), \bar{\phi}(\varphi)] + \mu_\beta \mathcal{A}(\mathcal{B}) \quad , \quad (2.16)$$

where  $\mu_\beta \equiv (n-1)/(4G) = (\beta - 2\pi)/(8\pi G)$  and  $\bar{g}(h)$ ,  $\bar{\phi}(\varphi)$  are corresponding classical solutions. The gravity action  $I[\bar{g}(h)]$  is given by (2.15). The total action  $I[\bar{g}(h), \bar{\phi}(\varphi)]$  in the r.h.s. of (2.16) is set on  $\bar{\mathcal{M}}$  and does not depend on  $\mathcal{B}$ . The partition function (2.13) can be written in the form

$$\mathcal{Z}(\beta, D_1, T) \sim e^{-nI[\bar{g}(h), \bar{\phi}(\varphi)]} \sum_{\mathcal{B}} e^{-\mu_\beta \mathcal{A}(\mathcal{B})} \quad . \quad (2.17)$$

The sum  $\sum_{\mathcal{B}}$  goes over surfaces embedable in  $\bar{\mathcal{M}}$ . Because  $\mu_\beta > 0$  the sum is dominated by the contribution from a surface  $\bar{\mathcal{B}}$  which has a least area in  $\bar{\mathcal{M}}$  for given boundary conditions. Therefore,

$$-\ln \mathcal{Z}(\beta, D_1, T) \simeq nI[\bar{g}(h), \bar{\phi}(\varphi)] + \mu_\beta \mathcal{A}(\bar{\mathcal{B}}) \quad . \quad (2.18)$$

By using (2.12) and (2.18) one gets the entanglement entropy in a quantum gravity theory for the region with boundary  $D_1$

$$S(D_1, T) \simeq \frac{\mathcal{A}(\bar{\mathcal{B}})}{4G} \quad . \quad (2.19)$$

The entropy is given in terms of the area of a surface separating the two regions with boundaries  $D_1$  and  $D_2$ . The formula is in agreement with the conjecture of [8]. As follows from (2.19), at the semiclassical level the entropy for the region with the boundary  $D_2$  has the same form,  $S(D_2, T) \simeq S(D_1, T)$ .

It is the presence of gravitational degrees of freedom in the functional integral (2.14) which results in (2.19). The reason why the entropy appears in the semiclassical approximation is explained by the fact that metric is an effective low-energy variable. One needs to know genuine degrees of freedom of the quantum gravity to write the entropy in a von Neumann form (2.2).

As technical remark, note that the partition function  $\mathcal{Z}(\beta, D_1, T)$  does not coincide with the Gibbons-Hawking path integral (2.7) in the limit  $\beta = 2\pi$ . The both functionals at the semiclassical level are determined by the same solution  $\bar{\mathcal{M}}$  to the Einstein equations. However, they differ by a multiplier  $\mathcal{N}$  present in  $\mathcal{Z}(\beta, D_1, T)$ . Formally  $\mathcal{N}$  counts a "number" of hypersurfaces embedable in  $\bar{\mathcal{M}}$  under given boundary conditions. This difference is not essential because one can remove  $\mathcal{N}$  by changing normalization of  $\mathcal{Z}(\beta, D_1, T)$ .

## 2.4 The Plateau problem

It is convenient to fix the position of  $\mathcal{B}$  with respect to some coordinate chart,  $X^\mu$ ,  $\mu = 1, \dots, d$ , on  $\bar{\mathcal{M}}$ . The embedding can be locally described in the parametric form

$$X^\mu = X^\mu(y^i) \quad , \quad (2.20)$$

where  $y^i$ ,  $i = 1, \dots, n = d - 2$  are some coordinates on  $\mathcal{B}$ . If  $\bar{g}_{\mu\nu}$  is the metric on  $\bar{\mathcal{M}}$  the metric induced on  $\mathcal{B}$  is  $\gamma_{ij} = \bar{g}_{\mu\nu} X^\mu_{,i} X^\nu_{,j}$ . The volume of  $\mathcal{B}$  is

$$\mathcal{A}(\mathcal{B}) = \int_{\mathcal{B}} d^n y \sqrt{\gamma} \quad , \quad (2.21)$$

where  $\gamma = \det \gamma_{ij}$ . Because hypersurfaces embedded in  $\bar{\mathcal{M}}$  differ by their position we look for local extrema of the volume functional with respect to the variation of coordinates  $X^\mu$  under the fixed metric  $\bar{g}_{\mu\nu}$  and fixed boundary,

$$\delta_X \mathcal{A}(\mathcal{B}) = 0 \quad . \quad (2.22)$$

Hypersurfaces with such a property are called *minimal* [20]. The entropy (2.19) is determined by a least volume minimal hypersurface, i.e. by a *global minimum* of the volume functional. Finding a minimal surface for a given boundary is called the *Plateau problem*. If the background metric  $\bar{g}_{\mu\nu}$  has a Lorentzian signature the hypersurfaces which obey condition (2.22) are called *extremal*.

The r.h.s. of (2.22) can be rewritten as

$$\delta_X \mathcal{A}(\mathcal{B}) = \int_{\mathcal{B}} d^n y \sqrt{\gamma} \gamma^{ij} X^\mu_{,i} X^\nu_{,j} \delta X_{\mu;\nu} \quad , \quad (2.23)$$

where  $\delta X_{\mu;\nu} = \bar{g}_{\mu\lambda} \nabla_\nu \delta X^\lambda$  and  $\nabla_\mu$  are the connections defined with respect to  $\bar{g}_{\mu\nu}$ . If  $l^\mu$  is a unit vector orthogonal to the boundary  $\partial\mathcal{B}$  and tangent to  $\mathcal{B}$  the boundary conditions imply that  $\delta X^\mu l_\mu = 0$  on  $\partial\mathcal{B}$ . From (2.22) and (2.23) one gets a Euclidean version of the Nambu-Goto equations

$$\Delta X^\mu + \Gamma^\mu_{\lambda\nu} X^\lambda_{,i} X^\nu_{,j} \gamma^{ij} = 0 \quad , \quad (2.24)$$

where  $\Delta$  is a Laplace operator on  $\mathcal{B}$  and  $\Gamma^\mu_{\lambda\nu}$  are the Christoffel coefficients for  $\bar{g}_{\mu\nu}$ .

The conditions for the minimal surface can be also written in another form. Let  $\delta X^\mu_\perp$  be a part of the vector  $\delta X^\mu$  which is orthogonal to  $\mathcal{B}$ . One can write  $\delta X^\mu_\perp = \alpha n^\mu + \beta p^\mu$ , where  $\alpha = n^\mu \delta X_\mu$ ,  $\beta = p^\mu \delta X_\mu$  and  $n^\mu$ ,  $p^\mu$  are two unit mutually orthogonal normals to  $\mathcal{B}$ . One finds

$$\delta_X \mathcal{A}(\mathcal{B}) = \int_{\mathcal{B}} d^n y \sqrt{\gamma} \gamma^{ij} X^\mu_{,i} X^\nu_{,j} \delta X_{\perp\mu;\nu} = \int_{\mathcal{B}} d^n y \sqrt{\gamma} (\alpha k_n + \beta k_p) \quad , \quad (2.25)$$

where  $k_n = \gamma^{\mu\nu} \nabla_\mu n_\nu$  and  $k_p = \gamma^{\mu\nu} \nabla_\mu p_\nu$  are traces of extrinsic curvatures of  $\mathcal{B}$  defined with the help of the projector  $\gamma_{\mu\nu} = g_{\mu\nu} - n_\mu n_\nu - p_\mu p_\nu$ . The contribution in (2.25) of the longitudinal part of the variation  $\delta X^\mu_\parallel$  (tangent to  $\mathcal{B}$ ) is a total divergence which vanishes due to the boundary conditions. Therefore, equations (2.23) are equivalent to

$$k_n = k_p = 0 \quad . \quad (2.26)$$

We complete this section with brief comments regarding the Plateau problem [20]. The theory of minimal surfaces appeared from consideration of soap films and it was known

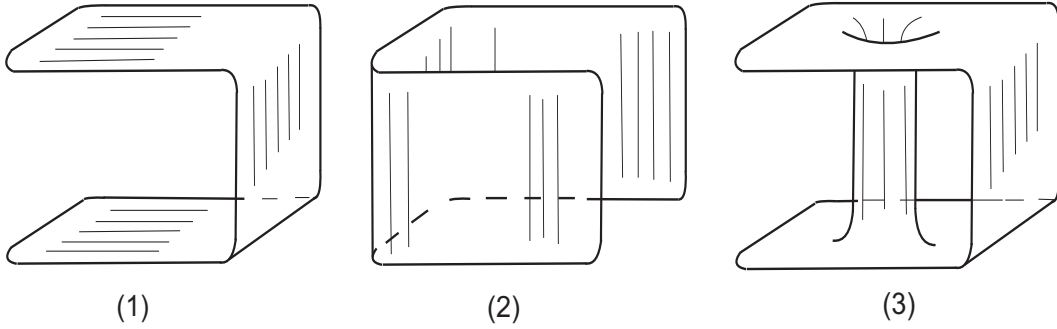


Figure 4: Some examples of minimal surfaces corresponding to a single contour in a flat space.

since 18-th century. The Plateau problem does not have a unique solution. Even in a flat space there may exist more than one minimal surface for a given boundary contour, see Fig. 4 (1,2,3). The solution is unique if the contour lies in a plane and does not have self-intersections. Note that in three-dimensions there may be contours which cannot be boundaries of minimal surfaces with a finite area. Such contours and their higher dimensional analogs are excluded from our analysis.

In general minimal surfaces can be represented as a combination of parts of smooth surfaces glued along their boundaries. These boundaries may appear as edges of the surfaces. The minimal surfaces may also have self-intersections and be non-orientable. Finally, the minimal surfaces may have a non-trivial topology, for example, they may have handles, see Fig. 4 (3). The least area minimal surface for a given boundary contour is not necessary a surface with a simplest topology. An interesting and a separate issue is the behavior (stability) of a minimal surface against small deformations.

Except soap films there are many other examples of minimal surfaces in Nature and human activities, e.g in the architecture. A DNA molecule, for instance, has a structure of double helix. The corresponding minimal surface is a helicoid, one of the four known minimal surfaces that can be embedded in a flat 3-dimensional space without self-intersections.

### 3 Suggestion about entanglement entropy in static space-times

The surface  $\mathcal{B}$  in the entropy formula (1.1) is embedded in a Riemannian manifold  $\bar{\mathcal{M}}$ . We suppose that  $\bar{\mathcal{M}}$  is a Euclidean section of some stationary Lorentzian space-time  $\bar{\mathcal{M}}^L$ . The metrics of the two spaces are related by the Wick rotation, i.e. by the complexification of the Killing time. If the metric is stationary but not static the Wick rotation is accompanied by an analytical continuation of some other parameters. For example, in an axially-symmetric space-time one goes to complex values of the angular velocity parameter. The same change of parameters has to be done in (1.1).

After the Wick rotation minimal co-dimension 2 hypersurfaces in  $\bar{\mathcal{M}}$  correspond to extremal hypersurfaces in  $\bar{\mathcal{M}}^L$ . In static Lorentzian space-times extremal hypersurfaces which end on a given slice  $\mathcal{S}$  of the boundary have an important geometrical property. They are minimal hypersurfaces located in a single constant time slice  $\Sigma_t$  such that  $\partial\Sigma_t =$

$\mathcal{S}$ . To prove this statement note that variations of the volume functional (2.22) in  $\bar{\mathcal{M}}^L$  can be written as  $\delta_X \mathcal{A} = \tilde{\delta}_X \mathcal{A} + \delta_t \mathcal{A}$  where  $\tilde{\delta}_X \mathcal{A}$  are caused by shifts of the surface along  $\Sigma_t$ . Variations  $\delta_t \mathcal{A}$  are generated by shifts  $\delta t$  along the time coordinate,  $\delta t = f(y)$  where  $f(y)$  is an arbitrary function of coordinates on  $\mathcal{B}$ . If  $\bar{\mathcal{B}}$  is minimal in  $\Sigma_t$  the variation of the volume becomes

$$\delta_X \mathcal{A}(\bar{\mathcal{B}}) = \delta_t \mathcal{A}(\bar{\mathcal{B}}) = \int_{\bar{\mathcal{B}}} d^n y \sqrt{\gamma} \gamma^{ij} X^\mu_{,i} X^\nu_{,j} f_{,\mu} \xi_\nu \quad . \quad (3.1)$$

We noticed here that  $\delta_t X^\mu = f(y) \xi^\mu$  and used the Killing identity  $\xi_{\mu;\nu} + \xi_{\nu;\mu} = 0$ . One concludes that a surface minimal in  $\Sigma_t$  is extremal in  $\bar{\mathcal{M}}^L$  if it is orthogonal to the Killing vector  $\xi$ . In static space-times the slice  $\Sigma_t$  is everywhere orthogonal to  $\xi$ . We come to the following statement based on the line of reasonings of Section 2.

Suggestion 1. *Suppose that: 1)  $\bar{\mathcal{M}}^L$  is a static Lorentzian space-time which is a solution to the Einstein equations; 2) constant Killing time  $t$  hypersurfaces,  $\Sigma_t$ , are the Cauchy hypersurfaces; 3) the boundary  $\mathcal{S}$  of  $\Sigma_t$  is simply connected. Consider a division of  $\Sigma_t$  into two regions with boundaries  $D_1$  and  $D_2$  such that  $D_1 \cup D_2 = \mathcal{S}$ . Suppose that  $\mathcal{B}$  is a minimal least volume codimension 1 hypersurface of  $\Sigma_t$  which ends at  $\mathcal{S}$  on the boundary of regions  $D_1$  and  $D_2$ .*

*Under these conditions the entanglement entropy in a quantum gravity theory associated to the partition of the system into the regions with external boundaries  $D_1$  and  $D_2$  is given by formula (1.1) where  $\mathcal{A}$  is the volume of  $\mathcal{B}$ . The formula holds in the leading order approximation when contributions of quantum fields are neglected.*

By the definition (see [21]) any non-space-like curve intersects the Cauchy hypersurface exactly once. The second assumption is needed because states in a quantum theory on a curved space-time are to be defined on a Cauchy hypersurface.

A possible extension of this suggestion to stationary but not static space-times will be discussed in section 8 after we examine additional requirements for the entropy formula.

## 4 Black holes

### 4.1 Cauchy surfaces with disconnected boundaries

Entanglement entropy in a quantum gravity (1.1) has the same form as the Bekenstein-Hawking entropy of a black hole. There are a number of arguments which indicate that entropy of black holes fits our approach as a particular case. The black hole geometries require a special consideration because they violate the condition that boundary of the Cauchy surface is connected.

To simplify the analysis we consider only the Schwarzschild solution

$$ds^2 = -B(r)dt^2 + \frac{dr^2}{B(r)} + r^2 d\Omega^2 \quad , \quad (4.1)$$

where  $B(r) = 1 - r_H/r$  and  $r > r_h$ . The metric describes the black hole geometry outside the horizon. The horizon is located at  $r = r_H$ . It is convenient to assume that the black hole is placed in spherical cavity with the radius  $r = r_0 > r_H$ .

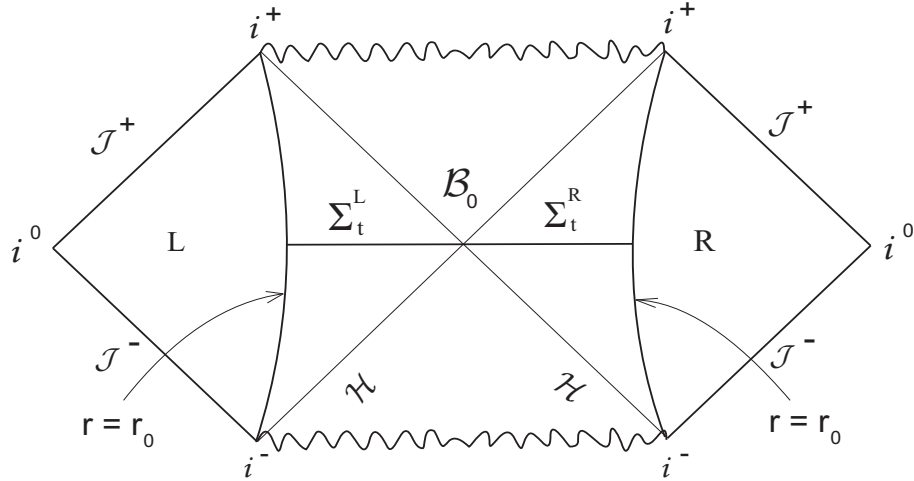


Figure 5: The Carter-Penrose diagram for a Schwarzschild black hole. The black hole in cavity corresponds to a part of the diagram between the lines  $r = r_0$ . Shown at the diagram is one of the Cauchy hypersurfaces  $\Sigma_t^L \cup \Sigma_t^R$ .

The Schwarzschild coordinates (4.1) are not complete. The entire geometry of an eternal black hole is shown on the Carter-Penrose diagram, Fig. 5. As is well known this geometry consists of two identical, left  $L$  and right  $R$ , regions separated by the horizons  $\mathcal{H}$ . The horizons intersect at a 2 sphere which is a bifurcation surface  $\mathcal{B}_0$ . Any space-like Cauchy hypersurface in this space-time stretches from the left world to the right and ends on the lines  $r = r_0$  or, in general, connects two asymptotically flat regions. All Cauchy hypersurfaces have two disconnected boundaries. The black hole geometry in the each region is static. A constant time Cauchy hypersurface  $t = \text{const}$  is called the Einstein-Rosen bridge. One of such surfaces is shown in Fig. 5. It consists of two parts,  $\Sigma_t^L$  and  $\Sigma_t^R$ , which join at  $\mathcal{B}_0$ . The Einstein-Rosen bridge has a wormhole topology. The boundary of the bridge is  $S^2 \cup S^2$ . We denote left and right pieces of the boundary as  $\mathcal{S}^L$  and  $\mathcal{S}^R$ , respectively.

As before, suppose that the system is divided into two parts with external boundaries  $D_1$  and  $D_2$  separated by a surface  $\mathcal{B}$ . A new possibility appears in case of black holes when  $D_1$  and  $D_2$  do not have common points, i.e.  $D_1 = \mathcal{S}^R$  and  $D_2 = \mathcal{S}^L$ . In this case the separating surface has to be a closed surface inside the Cauchy surface. We present the arguments that  $\mathcal{B} = \mathcal{B}_0$ .

The black hole in a cavity can be in a thermal equilibrium with its radiation [22]. Such a canonical ensemble is described by the partition function (2.7) with the boundary condition  $\mathcal{T} = S^2 \times S^1$ . The length of the circle  $S^1$  fixes the temperature of the ensemble, the size of the cavity is fixed by the radius of  $S^2$ . The semiclassical approximation to (2.7) is dominated by the contribution of the Gibbons-Hawking instanton [18]

$$ds^2 = B(r)d\tau^2 + \frac{dr^2}{B(r)} + r^2 d\Omega^2, \quad (4.2)$$

where  $\tau$  is periodic coordinate,  $0 \leq \tau \leq \beta_H$ , and  $\beta_H = 4\pi r_H$  is the inverse Hawking temperature. The temperature of the ensemble is  $T = 1/(\beta_H \sqrt{B(r_0)})$ .

The black hole is in a pure state which is called the Hartle-Hawking vacuum. A

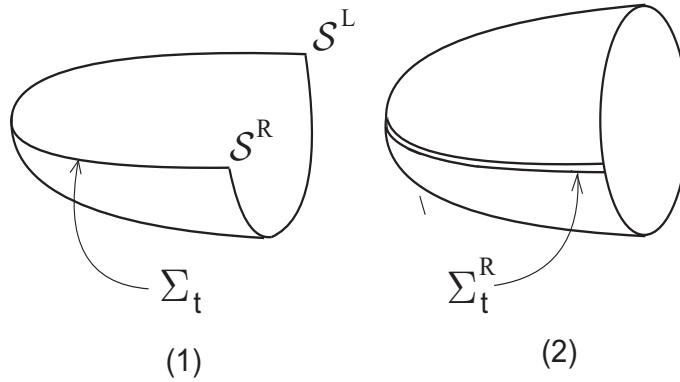


Figure 6: demonstrates the form of trajectories in the path integral representation for the wave function  $\Psi$  of the black hole and the density matrix  $\hat{\rho}$ , figures (1) and (2), respectively. Fig. (1) shows half of the Gibbons-Hawking instanton; (2) shows the instanton with a cut.

thermal nature of states in  $\Sigma_t^R$  appears when the region  $\Sigma_t^L$  is unobservable. Such point of view was suggested long ago [23] in the framework of a thermo-field dynamics.

A more general concept, a wave-function of a black hole,  $\Psi(\varphi)$ , was introduced in [24]. Semiclassically  $\Psi(\varphi)$  generates the Hartle-Hawking vacuum. The arguments  $\varphi$  of  $\Psi$  (a configuration space of the theory) are 3-geometries,  $\Sigma$ , and matter fields on  $\Sigma$ . Spaces  $\Sigma$  have the wormhole topology  $R \times S^2$ . Their boundary consists of two disconnected pieces,  $\mathcal{S}^L$  and  $\mathcal{S}^R$ , or two asymptotically flat regions, in general.  $\Psi(\varphi)$  is defined by a path integral like (2.7). The trajectories here are compact 4-geometries with the boundary  $\Sigma \cup (S^2 \times I)$  where  $I$  is an interval of the length  $1/(2T)$ . An example of such a geometry is shown on Fig. 6 (1). It is a half of the Gibbons-Hawking instanton obtained by cutting the instanton along the bridge ( $\tau = 0$  and  $\tau = \beta_H/2$ ).

Consider a Gibbons-Hawking instanton with a cut along  $\Sigma_t^R$ , Fig. 6 (2). Let us take  $n$  identical copies of the given space and glue them along the cuts. We obtain a 4-geometry,  $\bar{\mathcal{M}}_n$ , which has conical singularities located at the horizon  $r = r_h$ . Locally  $\bar{\mathcal{M}}_n$  looks as (4.2) but has the periodicity  $0 \leq \tau \leq n\beta_H$ . The geometry above can be used to compute the entanglement entropy of field excitations located on  $\Sigma_t^R$  with the help of formula (2.6). To this aim one has to use the partition function (2.5) on  $\bar{\mathcal{M}}_n$ . The instanton with a cut corresponds to the density matrix  $\hat{\rho}$  of fields in  $\Sigma_t^R$ .

To compute the entanglement entropy in a quantum gravity theory we, like the authors of [24], allow quantum fluctuations of the geometry. Both the Gibbons-Hawking instanton and position of the separating surface are to appear semiclassically when one uses the approach of section 2 with boundary conditions adopted to black holes.

We suggest to this aim that the entanglement entropy for the region with the boundary  $D_1 = \mathcal{S}^R$  and a black hole inside is determined by a partition function  $\mathcal{Z}(\beta, D_1, T)$ , see (2.12). The partition function is represented in terms of the integrals (2.13), (2.14). The path integral (2.14) it is taken over a class of Riemannian 4-geometries,  $\{\mathcal{M}_n\}_{\mathcal{B}}$  which have the topology of  $\bar{\mathcal{M}}_n$ . The boundary of these spaces coincides with the boundary of  $\bar{\mathcal{M}}_n$ . The class  $\{\mathcal{M}_n\}_{\mathcal{B}}$  is specified by the presence of an internal co-dimension 2 hypersurface  $\mathcal{B}$  where the spaces have conical singularities  $\mathcal{C}_\beta \times \mathcal{B}$ ,  $\beta = 2\pi n$ .

In the semiclassical approximation the integral (2.14) is determined by a trajectory which locally has the same geometry as (4.2). The sum (2.13) picks up a class where the

surface  $\mathcal{B}$  has a least area on (4.2). Because (4.2) is a static space it is enough to find a least area surface  $\bar{\mathcal{B}}$  lying in the Einstein-Rosen bridge. The spherical symmetry of  $\Sigma_t$  implies that  $\bar{\mathcal{B}}$  has a constant radius  $r$  of the minimal value  $r = r_H$ . Thus, the entanglement entropy is given by (1.1) and coincides with the Bekenstein-Hawking entropy in accord with the interpretation of [9]–[11]<sup>2</sup>.

The surface  $r = r_H$  also solves equations (2.25). The Killing field  $\partial_\tau$  has fixed points at  $r = r_H$  and, as a result of this symmetry extrinsic curvatures of  $\bar{\mathcal{B}}$  vanish. Therefore,  $\bar{\mathcal{B}}$  is minimal according to (2.26). The same property holds for the bifurcation surface of the horizon of rotating black holes.

A spherical form of a minimal surface in a black hole space-time is supported by the strong gravitational field near the horizon. Closed minimal surfaces with the topology of a sphere cannot exist in a flat space. An analog of a closed minimal surface in a 3D flat space is a soap bubble. The bubble has a constant non-vanishing mean curvature whose value is determined by the surface tension and by the difference of pressures inside and outside.

## 4.2 Entanglement in the presence of a black hole

In case of black holes the choice of the region where the states of the system are traced out can be more non-trivial and it may not coincide with the division of the system by the horizon. Moreover, the separating surface may consist of several disconnected elements including the black hole horizon as one of them. Several possibilities are shown in Fig. 7. We suppose here that the boundary  $\mathcal{S}^R$  of the right half of the Einstein-Rosen bridge is divided into two parts,  $\mathcal{S}_1^R$  and  $\mathcal{S}_2^R$ . (The left part of boundary is  $\mathcal{S}^L$ .) A least area surface inside the bridge starts at the boundary between  $\mathcal{S}_1^R$  and  $\mathcal{S}_2^R$ . Locations of this surface are denoted as  $\mathcal{B}_1$  or  $\mathcal{B}_2$ , depending on its position with respect to the horizon  $\mathcal{B}_0$ . Arguments given in section 2.3 indicate that among different options one has to choose the one with the least total area.

In quantum gravity the entanglement is determined by the division of the boundary of the system into parts  $D_1$  and  $D_2$ . As earlier, we assume that the degrees of freedom in a region with the boundary  $D_2$  are integrated out. Dashed parts on Fig. 7 show locations of this region on the Einstein-Rosen bridge for different choices of the boundary.

In cases (1),(2) the boundary is  $D_1 = \mathcal{S}_1^R$ ,  $D_2 = \mathcal{S}^L \cup \mathcal{S}_2^R$ . The first possibility, (1), is realized when  $\mathcal{A}(\mathcal{B}_1) < \mathcal{A}(\mathcal{B}_2) + \mathcal{A}(\mathcal{B}_0)$ , the second, (2), in the opposite case. The third option is  $D_1 = \mathcal{S}^L \cup \mathcal{S}_2^R$ ,  $D_2 = \mathcal{S}_1^R$ , Fig. 7, (3). It is dual to the case (1) and is realized when states can be measured both on the right and on the left parts of the bridge, see Fig. 5. The options (1) and (3) have the same entanglement entropy,  $\mathcal{A}(\mathcal{B}_1)/(4G)$ , because the system is in a pure state, the Hartle-Hawking vacuum.

When states inside the black hole cannot be observed, instead of (3) one has to consider another definition of the boundary. It is  $D_1 = \mathcal{S}_2^R$ ,  $D_2 = \mathcal{S}^L \cup \mathcal{S}_1^R$ , see Fig. 7, (4),(5). The corresponding entropy is either  $(\mathcal{A}(\mathcal{B}_1) + \mathcal{A}(\mathcal{B}_0))/(4G)$  or  $\mathcal{A}(\mathcal{B}_2)/(4G)$ . In the latter case  $\mathcal{A}(\mathcal{B}_2) < \mathcal{A}(\mathcal{B}_1) + \mathcal{A}(\mathcal{B}_0)$ . In both situations the entropy differs from the entropy in (1) because observers outside the horizon perceive the black hole in a thermal state. We

---

<sup>2</sup>A similar interpretation of the Bekenstein-Hawking entropy on the base of minimal surfaces is given in [25] for black holes on a brane. Different aspects of entanglement entropy as a source of black hole entropy were discussed recently in [26]–[29].

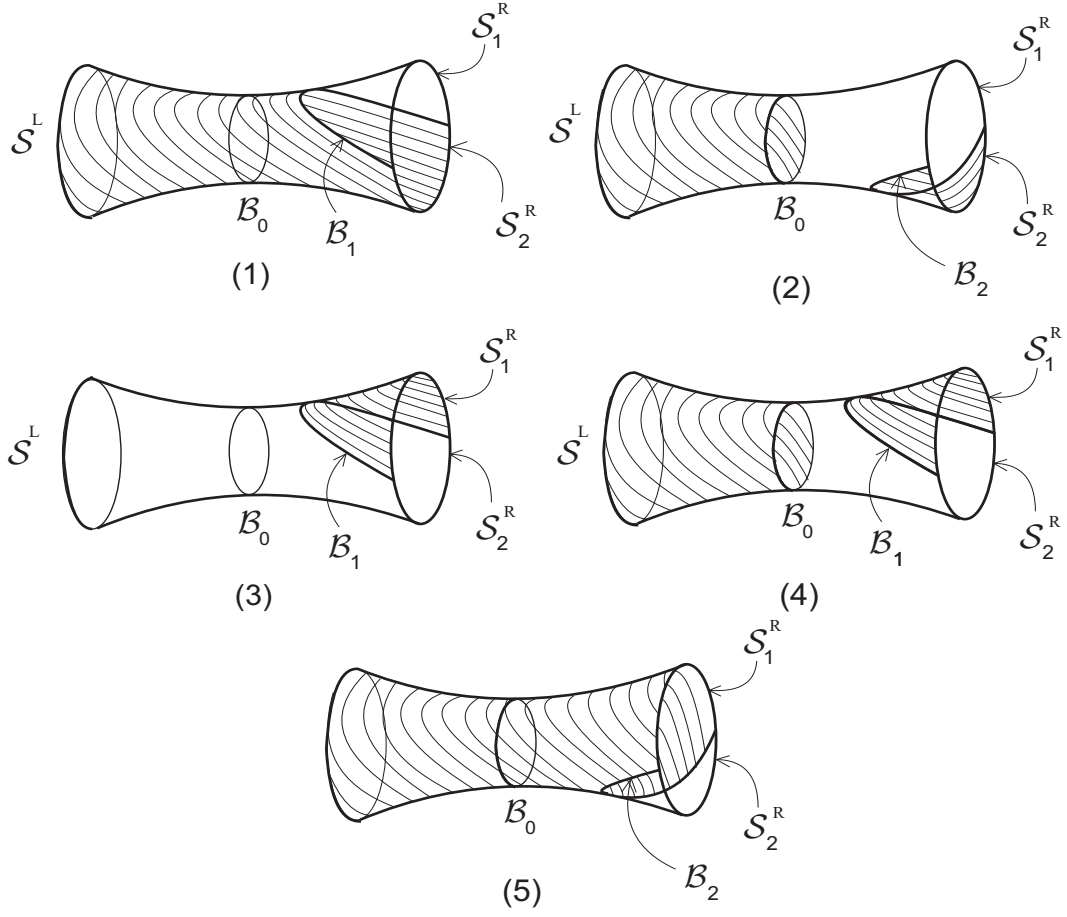


Figure 7: Shown by dashed regions are different locations of states on the Einstein-Rosen bridge which are traced out. The boundary of these regions is  $D_2$ , the boundary of their completion is  $D_1$ . The cases (1),(2) correspond to the choice  $D_1 = S_1^R$ ,  $D_2 = S^L \cup S_2^R$ , the case (3) is  $D_1 = S^L \cup S_2^R$ ,  $D_2 = S_1^R$ , and the cases (4),(5) are  $D_1 = S_2^R$ ,  $D_2 = S^L \cup S_1^R$ .

come to the following statement regarding entanglement entropy in the presence of black holes.

Suggestion 2: suppose that: 1)  $\bar{\mathcal{M}}^L$  is a Lorentzian space-time which is a static black hole solution to the Einstein equations; 2)  $\Sigma_t$  is a part of a constant-Killing-time hypersurface in  $\bar{\mathcal{M}}^L$  restricted by the black hole horizon  $\mathcal{B}_0$  and by an external boundary  $\mathcal{S}$ . Consider a division of  $\Sigma_t$  into two regions with boundaries  $\tilde{D}_1$  and  $\tilde{D}_2$  such that  $\tilde{D}_1 \cup \tilde{D}_2 = \mathcal{S}$ . Let  $\mathcal{B}_1$  and  $\mathcal{B}_2$  be codimension 1 hypersurfaces in  $\Sigma_t$  which end on the boundary between  $\tilde{D}_1$  and  $\tilde{D}_2$  and have least volumes among the hypersurfaces homologous to  $\tilde{D}_1$  or  $\tilde{D}_2$ , respectively, see Fig. 8.

The entanglement entropy in a quantum gravity theory associated to the loss of the information about the states in the region with boundary  $\tilde{D}_2$  is

$$S(\tilde{D}_1, T) = \frac{\mathcal{A}(\mathcal{B}_1)}{4G} \quad , \quad (4.3)$$

if  $\mathcal{A}(\mathcal{B}_1) < \mathcal{A}(\mathcal{B}_2) + \mathcal{A}(\mathcal{B}_0)$ , and

$$S(\tilde{D}_1, T) = \frac{\mathcal{A}(\mathcal{B}_2)}{4G} + \frac{\mathcal{A}(\mathcal{B}_0)}{4G} \quad (4.4)$$



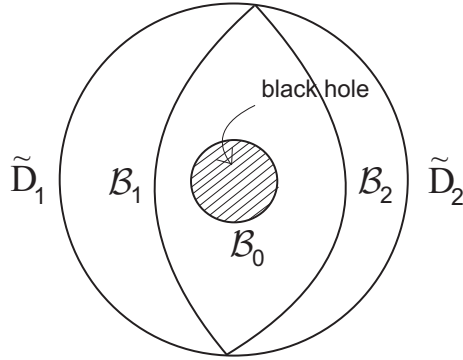


Figure 8: Separating surfaces in the presence of a black hole.

in the opposite case. The formulae hold in the leading order approximation when contributions of quantum fields are neglected. If  $\tilde{D}_2 = \emptyset$  the entropy is the Bekenstein-Hawking entropy of the black hole. Parameter  $T$  in (4.3), (4.4) is a temperature of the black hole.

## 5 Entropy inequalities

The suggestions made in Sections 2 and 4 cannot be rigorously proved in an effective gravity theory because a real microscopical content of the underlying theory is not known. The situation is similar to the problem of the Bekenstein-Hawking entropy  $S^{BH}$ . The origin of the entropy of a black hole can be understood only in a genuine quantum gravity theory. The fact that  $S^{BH}$  has properties of a physical entropy follows from different thought experiments which show, for example, that a generalized version of the second law of thermodynamics holds for black holes. In a similar way, one can check that suggestions 1 and 2 for the entanglement entropy are physically and mathematically consistent. The first step is to see whether they agree with basic properties of the von Neumann entropy.

In the context of the holographic representation of the entanglement entropy in theories with AdS duals, it was proposed [15] to use the subadditivity property as a test for definition of the entropy. Later Headrick and Takayanagi [16] gave a simple geometrical proof that the entropy in such theories possesses this property. The proof of [16] was based on the relation between the entropy and minimal surfaces. That is why it is also applicable under our conditions.

Consider a set of states,  $a$ , in a Hilbert space of some quantum system. We assume that  $a$  corresponds to states located in some region of the coordinate space. Let us define the reduced density matrix  $\hat{\rho}_a = \text{Tr}_{\bar{a}} \hat{\rho}$  where  $\hat{\rho}$  is the density matrix of the entire system. The trace  $\text{Tr}_{\bar{a}}$  is taken over the Hilbert space except elements  $a$ . The entanglement entropy related to the given set is  $S_a = -\text{Tr} \hat{\rho}_a \ln \hat{\rho}_a$ . For the entropies of two sets shown on Fig. 9, (1) one can write the following inequalities:

$$S_1 + S_2 \geq S_{1 \cup 2} + S_{1 \cap 2} \quad , \quad (5.1)$$

$$|S_1 - S_2| \leq S_{1 \cup 2} \quad , \quad (5.2)$$

where  $S_{1 \cup 2}$  and  $S_{1 \cap 2}$  are the entropies for the union or intersection of the sets 1 and 2. Formulae above are based on the concavity of the function  $-x \ln x$  which is present in

the von Neumann definition of the entropy [30]. Equation (5.1) is known as the strong subadditivity property. The Araki-Lieb inequality (5.2) holds when the Hilbert space of the system consists only of the two components, 1 and 2.

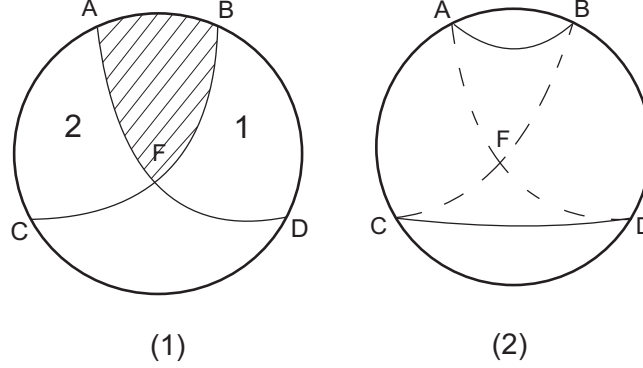


Figure 9: Fig. (1) demonstrates a system with intersecting subsets (regions), 1 and 2. Boundaries of the regions are shown as lines  $(A, D)$  and  $(B, C)$ . The domain of the intersection is dashed. The minimal surfaces are the internal lines connecting pairs of points,  $A$  with  $D$  and  $B$  with  $C$ . Another set of minimal surfaces for the same configuration of the boundaries are the solid internal lines on Fig. (2). They connect  $A$  with  $B$  and  $C$  with  $D$ .

The sketch of the proof of (5.1) is the following<sup>3</sup>. If the system does not have a black hole inside then entropies  $S_1$  and  $S_2$  are given by the areas of least area surfaces shown on Fig. 9, (1), as internal lines connecting, respectively, pairs of points,  $A$  with  $D$  and  $C$  with  $B$ . The entropies  $S_{1 \cup 2}$  and  $S_{1 \cap 2}$  are given by least area surfaces shown as internal lines connecting, respectively,  $C$  with  $D$  and  $A$  with  $B$ , see Fig. 9, (2). Let us denote as  $\mathcal{A}_{XY}$  the area of a least minimal surface corresponding to an internal line on Fig. 9 connecting points  $X$  and  $Y$ . Then

$$S_1 + S_2 = \mathcal{A}_{AD} + \mathcal{A}_{BC} = \mathcal{A}_{AF} + \mathcal{A}_{FD} + \mathcal{A}_{BF} + \mathcal{A}_{FC} =$$

$$(\mathcal{A}_{AF} + \mathcal{A}_{BF}) + (\mathcal{A}_{FD} + \mathcal{A}_{FC}) \geq \mathcal{A}_{AB} + \mathcal{A}_{DC} = S_{1 \cup 2} + S_{1 \cap 2} \quad . \quad (5.3)$$

Let us discuss now the Araki-Lieb inequality (5.2). If the state of the system is pure,  $S_{1 \cup 2} = 0$ . In this case (5.2) is saturated because of the known symmetry property,  $S_1 = S_2$ , of the entanglement entropy in pure states. This property is trivially satisfied if there are no black holes: the least area hypersurface separating the system is uniquely determined because the boundary of the Cauchy surface is simply connected.

In case of black holes the situation is more complicated because there may be two types of the least area surfaces homologous to different parts of the boundary, see Fig. 7. Black hole is perceived by an external observer as a mixed state whose entropy is given by the Bekenstein-Hawking formula. Thus, (5.2) should be written as

$$|S_1 - S_2| \leq S = S_0 \quad , \quad (5.4)$$

where we used notations  $S_1 = S(\tilde{D}_1, T)$ ,  $S_2 = S(\tilde{D}_2, T)$ ,  $S_0 = \mathcal{A}(\mathcal{B}_0)$ . Let us show that (5.4) follows from the suggestion 2. Denote  $\mathcal{A}_k = \mathcal{A}(\mathcal{B}_k)$ , where  $k = 0, 1, 2$ . Suppose

<sup>3</sup>For simplicity in this section we put  $4G = 1$ .

that  $\mathcal{A}_2 > \mathcal{A}_0$ , and consider different situations. If  $\mathcal{A}_1 < \mathcal{A}_2 - \mathcal{A}_0$  one easily concludes that  $S_1 = \mathcal{A}_1$ ,  $S_2 = \mathcal{A}_1 + \mathcal{A}_0$ , and the Araki-Lieb inequality is saturated. The same happens when  $\mathcal{A}_1 > \mathcal{A}_2 + \mathcal{A}_0$  because in this case  $S_1 = \mathcal{A}_2 + \mathcal{A}_0$ ,  $S_2 = \mathcal{A}_2$ . Finally, if  $\mathcal{A}_2 - \mathcal{A}_0 < \mathcal{A}_1 < \mathcal{A}_2 + \mathcal{A}_0$  the entropies are  $S_1 = \mathcal{A}_1$ ,  $S_2 = \mathcal{A}_2$ . Then

$$S_2 - S_1 = \mathcal{A}_2 - \mathcal{A}_1 < \mathcal{A}_0 = S_0 \quad \text{for } S_2 > S_1 \quad ,$$

$$S_1 - S_2 = \mathcal{A}_1 - \mathcal{A}_2 < \mathcal{A}_0 = S_0 \quad \text{for } S_1 > S_2 \quad .$$

The case  $\mathcal{A}_2 < \mathcal{A}_0$  is treated analogously.

Extension of these results to stationary but not static space-times will be discussed in section 8.

## 6 Variational formulae

In general, in many-body systems the entanglement entropy cannot be expressed solely in terms of macroscopical parameters. Additional information about the microscopical structure is needed. That is why, in contrast with the first law for the thermodynamical entropy, there are no universal variational formulae for the entanglement entropy of ordinary quantum systems. According to our suggestions the entanglement entropy  $S$  in a quantum gravity is a geometrical quantity.

We give examples how to express changes of  $S$  solely in terms of macroscopical parameters. In Minkowsky space the minimal surface which divides the space into two parts is a plane. If one brings in a point-like particle with mass  $M$  the shape of the surface slightly changes under the influence of the gravitational field of the particle. The space-time metric in the weak field approximation is

$$ds^2 = - \left(1 - \frac{2MG}{R}\right) dt^2 + \left(1 + \frac{2MG}{R}\right) (dx^2 + dy^2 + dz^2) \quad , \quad (6.1)$$

where  $R = \sqrt{x^2 + y^2 + z^2}$  and the particle is located at  $R = 0$ . It is not difficult to see that the area of the minimal surface in the weak field approximation is given by the integral

$$\mathcal{A} = \int dx dy \left(1 + \frac{2MG}{\sqrt{x^2 + y^2 + z^2}}\right) \quad , \quad (6.2)$$

where  $z$  is a coordinate of the plane in the absence of the particle. To cutoff the integral (6.2) one may assume that the surface has a large but finite size. The variation of the area of the surface under the change of its position with respect to the particle does not depend on the cutoff

$$\delta \mathcal{A} = -\delta z \int dx dy \frac{2MGz}{(x^2 + y^2 + z^2)^{3/2}} = -4\pi MG \delta z \quad . \quad (6.3)$$

The area decreases if the distance  $z$  increases. If one restores all dimensional constants the variational formula for the entanglement entropy takes the following form

$$\delta S = -\pi \frac{Mc}{\hbar} \delta z \quad . \quad (6.4)$$

For a particle of the mass 1 *g* the change of the entropy under the shift 1 *cm* is an enormous quantity,  $2.7 \cdot 10^{37}$ . For an elementary particle (6.4) can be written as

$$\delta S = -\pi \frac{\delta z}{\lambda} \quad , \quad (6.5)$$

where  $\lambda = \hbar/(Mc)$  is the Compton wave length of the particle. Thus, if the particle shifts to the distance comparable to its Compton wave length the entropy change is of the order of unity.

Another example where the variation of the entropy can be easily established is the space-time around a cosmic string

$$ds^2 = -dt^2 + (1 - 4\mu G)^2 \rho^2 d\varphi^2 + d\rho^2 + dz^2 \quad , \quad (6.6)$$

where  $0 \leq \varphi < 2\pi$ . The string is located at  $\rho = 0$  and has a tension  $\mu$ . The gravitational effects around the string result in a deficit of the polar angle around its axis. A minimal surface which is parallel to the string is a plane slightly distorted by the string gravity. Suppose that in the Cartesian coordinates ( $x = \rho \cos \varphi$ ,  $y = \rho \sin \varphi$ ) the position of the plane is  $x = l$ . Then in the weak field approximation,  $\mu G \ll 1$ , the area of the minimal surface per unit length is

$$a = \int dy \left( 1 - 4\mu G \frac{l^2}{l^2 + y^2} \right) \quad . \quad (6.7)$$

The parameter  $l$  is a distance from the string to the minimal surface. Note that the deviation from the area of the plane per unit length is finite,  $-4\pi\mu Gl$ . Change of the entropy (per unit length) under the shift of the string  $\delta l$  directed outward of the minimal surface is

$$\delta s = -\pi\mu\delta l \quad . \quad (6.8)$$

In Planck units  $\mu \sim 10^{-6}$  for GUT strings ( $10^{16}$  *GeV*) and  $\mu \sim 10^{-34}$  for electroweak strings ( $100$  *GeV*). One can consider deformations of the strings when a portion of the string of a size  $z$  moves toward or outward of the plane to a distance  $\delta l$ . Changes of the entropy caused by such deformations can be estimated as  $\delta S \sim z\delta s \sim -\mu z\delta l$ . It is easy to see that  $\delta S$  is of the order of unity when both  $z$  and  $\delta l$  are determined by the scales of the corresponding theory, ( $z \sim \delta l \sim 10^{-30}$  *cm* for GUT strings, or  $z \sim \delta l \sim 10^{-16}$  *cm* for electroweak strings). If  $z$  or  $\delta l$  are macroscopical scales the entropy changes are very large.

## 7 Quantum corrections

In the Gibbons-Hawking approach the low-energy fields  $\phi$  yield quantum corrections to the semiclassical approximation (2.10). Quantum effects change the classical solution  $\bar{g}_{\mu\nu}$  because of a stress-energy tensor of the quantum matter in the r.h.s. of the Einstein equations (2.11). As a result, deformations of the minimal surface  $\mathcal{B}$  caused by the back-reaction change of the entanglement entropy (1.1).

There may be another source of quantum corrections to the entropy. Entanglement of the low-energy fields separated by the surface  $\mathcal{B}$  is quantified by its own entropy. We denote this entropy  $S_q(\mathcal{B})$ . In an effective theory  $S_q(\mathcal{B})$  is an ultraviolet divergent quantity

and leading divergences of  $S_q(\mathcal{B})$  are proportional to the area  $\mathcal{A}(\mathcal{B})$  [9], [10]<sup>4</sup>. Thus, the "gravitational" entropy,  $\mathcal{A}(\mathcal{B})/(4G)$ , and  $S_q(\mathcal{B})$  may be related. We make the following statement.

*Suggestion 3: In a static space-time entanglement entropy  $S_q(\mathcal{B})$  of the low-energy matter fields divided by a minimal hypersurface  $\mathcal{B}$  contributes to a quantum correction to the entropy (1.1). The leading ultraviolet divergences of  $S_q(\mathcal{B})$  are removed in the course of a standard renormalization of the gravity coupling  $G$ .*

The idea of renormalizing the divergences in the entanglement entropy was first formulated for fields around a black hole [32],[33] (for a review, see [34]). We emphasize that reasonings applicable to black hole horizons can be extended to minimal hypersurfaces in a constant-time slice.

Subleading divergences in the effective action require terms which depend quadratically on components of the curvature tensor. The form of the entropy (1.1) in the presence of  $R^2$ -terms is modified. Once this modification is taken into account, we suppose that subleading divergences of  $S_q(\mathcal{B})$  are also eliminated in the course of renormalization of the corresponding couplings at  $R^2$ -terms.

Suggestion 3 gives a prescription how to do computations which take into account entanglement entropy of matter fields. Its consequences and applications will be considered elsewhere.

## 8 Concluding remarks

We presented a line of reasonings that the entanglement entropy  $S$  in a quantum gravity theory is defined by macroscopical parameters of the system. Like a thermodynamical entropy,  $S$  does not require knowing a genuine microscopical structure. In the leading order  $S$  is given by the Bekenstein-Hawking formula (1.1) in terms of the volume of a co-dimension 2 hypersurface  $\mathcal{B}$ . For given boundary conditions  $\mathcal{B}$  must be a least volume hypersurface embedded in a Euclidean section  $\bar{\mathcal{M}}$  of the corresponding Lorenzian solution. Our reasonings are based on the assumption that gravity is an emergent phenomenon and we make use of the Gibbons-Hawking path integral as a low-energy definition of the partition function.

In static space-times  $\mathcal{B}$  can be interpreted as a real hypersurface on a constant time slice. Thus, despite of fluctuations of the geometry a spatial division of degrees of freedom in a quantum gravity theory can have a meaning in a semiclassical approximation provided that the separating surface has a least volume in the slice.

The idea that the concept of a "gravitational" entropy can be extended to co-dimension 2 hypersurfaces is not new. In [35] Jacobson applied the Bekenstein-Hawking formula to small space-like 2-surfaces in arbitrary 4D space-times. The past directed null normal congruences of the surfaces ("local Rindler horizons") were supposed to have vanishing expansion and shear at a point. In [36] an attempt has been made to extend this construction to arbitrary 2-surfaces with a finite area. What differs our approach from the above suggestions is that we directly relate the entropy to quantum entanglement and on this base derive dynamical equations for the surface.

---

<sup>4</sup>For the discussion of the divergences of the entanglement entropy for a system with a boundary in Minkowsky space-time, see [8], [31] and references therein.

In static space-times properties of the entanglement entropy  $S$  are relatively simple. Although the underlying degrees of freedom are not known one can demonstrate here that  $S$  satisfies a subadditivity property, which is a necessary requirement. This happens because  $\mathcal{B}$  is defined as a least volume hypersurface.

The next step would be to define the entropy in case of stationary but not static space-times. Our approach indicates that  $S$  is still given by (1.1) where  $\mathcal{B}$  is a minimal least volume co-dimension 2 hypersurface on a Euclidean section  $\bar{\mathcal{M}}$  of a corresponding physical space-time  $\bar{\mathcal{M}}^L$ . The physical value of the entropy should be attained after analytical continuation. After the Wick rotation minimal co-dimension 2 hypersurfaces in  $\bar{\mathcal{M}}$  correspond to extremal hypersurfaces in  $\bar{\mathcal{M}}^L$ .

If  $\bar{\mathcal{M}}^L$  is stationary but not static an extremal surface  $\mathcal{B}$  does not lie in a constant-time slice unless  $\mathcal{B}$  is orthogonal to the Killing field generating time translations, see section 3. Thus, in general different surfaces  $\mathcal{B}$  may separate the system in different space-like slices.

One can still try to give a definition of the entropy which is consistent with the basic features such as the subbadditivity property. It is natural to guess that in stationary space-times  $\mathcal{B}$  in (1.1) should be an extremal hypersurface with a least volume for given boundary conditions. In certain examples such surfaces can be obtained from the corresponding least volume surfaces in Euclidean theory. If this is the case one can prove the subbadditivity property along the lines of section 5. The proof in section 5 does not require the surfaces to be in a single slice.

Generalization of our construction to time-dependent geometries should be the subject of a separate discussion. Some proposals in the context of the holographic formula for the entanglement entropy in conformal theories are discussed in [37].

## Acknowledgment

I am grateful N.A. Tyurin and A.I. Zelnikov for very helpful discussions. The author thanks the Asia Pacific Centre for Theoretical Physics for the support during the Focus Program on New Frontiers in Black Hole Physics (Pohang, 2007). This work was also supported by the Scientific School Grant N 5332.2006.2.

## References

- [1] A.D. Sakharov, Sov. Phys. Dokl. **12** (1968) 1040.
- [2] G.E. Volovik, *The Universe in a Helium Droplet*, Clarendon Press, Oxford (2003).
- [3] T. Padmanabhan, *Gravity as an Emergent Phenomenon: a Conceptual Description*, e-Print: arXiv:0706.1654 [gr-qc].
- [4] G.E. Volovik and A.I. Zelnikov, *Universal Temperature Corrections to the Free Energy for the Gravitational Field*, e-Print: gr-qc/0309066.
- [5] C.P. Burgess, Liv. Rev. Rel. 7:5 (2004), e-Print: gr-qc/0311082; C.P. Burgess, *Introduction to Effective Field Theories*, e-Print: hep-th/0701053.
- [6] L. Amico, R. Fazio, A. Osterloch, V. Vedral, *Entanglement in Many-Body Systems*, e-Print: quant-ph/0703044.
- [7] J. Preskill, J. Mod. Opt. **47** (2000) 127, e-Print: quant-ph/9904022.
- [8] D.V. Fursaev, Phys. Rev. **D73** (2006) 124025, e-Print: hep-th/0602134.
- [9] M. Srednicki, Phys. Rev. Lett. **71** (1993) 666, e-Print: hep-th/9303048.
- [10] L. Bombelli, R.K. Koul, J. Lee and R.D. Sorkin, Phys. Rev. **D34** (1986) 373.
- [11] V. Frolov and I. Novikov, Phys. Rev. **D48** (1993) 4545.
- [12] S. Ryu and T. Takayanagi, Phys. Rev. Lett. **96** (2006) 181602, e-Print: hep-th/0603001.
- [13] S. Ryu and T. Takayanagi, JHEP **0608** (2006) 045, e-Print: hep-th/0605073.
- [14] D.V. Fursaev, JHEP **0609** (2006) 018, e-Print: hep-th/0606184.
- [15] T. Hirata and T. Takayanagi, JHEP **02** (2007) 042 e-Print: hep-th/0608213.
- [16] M. Headrick and T. Takayanagi, *A holographic proof of the strong subadditivity of entanglement entropy*, e-Print: arXiv:0704.3719 [hep-th].
- [17] G.W. Gibbons and S.W. Hawking, Phys. Rev. **D15** (1977) 2752.
- [18] S.W. Hawking, In: *General Relativity: An Einstein Centenary Survey*. (eds. S.W. Hawking and W. Israel), Cambridge Univ.Press, Cambridge, 1979.
- [19] D.V. Fursaev and S.N. Solodukhin, Phys. Rev. **D52** (1995) 2133, e-Print: hep-th/9501127.
- [20] Dao Trong Thi, A.T. Fomenko, *Minimal Surfaces and the Plateau Problem*, Nauka, Moscow, 1987 (in Russian).
- [21] S.W. Hawking and G.F.R. Ellis, *The Large Scale Structure of Space-Time*, Cambridge University Press, 1973.

- [22] J.W. York, Phys. Rev. **D33** (1986) 2092.
- [23] W. Israel, Phys. Lett. **57A** (1976) 107.
- [24] A.O. Barvinsky, V.P. Frolov, and A.I. Zelnikov, Phys. Rev. **D51** (1995) 1741.
- [25] R. Emparan, JHEP **0606** (2006) 012, e-Print: hep-th/0603081.
- [26] S.N. Solodukhin, Phys. Rev. Lett. **97** (2006) 201601, e-Print: hep-th/0606205.
- [27] M. Cadoni, Phys. Lett. **B653** (2007) 434, e-Print: arXiv:0704.0140 [hep-th], *Induced gravity and entanglement entropy of 2D black holes* e-Print: arXiv:0709.0163 [hep-th].
- [28] R. Brustein, M.B. Einhorn, and A. Yaron, JHEP **0601** (2006) 098, e-Print: hep-th/0508217.
- [29] T. Azeyanagi, T. Nishioka, T. Takayanagi, *Near extremal black hole entropy as entanglement entropy via  $AdS_2/CFT_1$* , e-Print: arXiv:0710.2956 [hep-th].
- [30] M.A. Nielsen and I.L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, 2000.
- [31] H. Casini and M. Huerta, *Analytic results on the geometric entropy for free fields*, e-Print: arXiv:0707.1300 [hep-th]; H. Casini and M. Huerta, Nucl. Phys. **B764** (2007) 183. e-Print: hep-th/0606256.
- [32] C. Callan and F. Wilczek, *On Geometric Entropy*, Phys. Lett **B333** (1994) 55.
- [33] L. Susskind and J. Uglum, *Black Hole Entropy In Canonical Quantum Gravity And Superstring Theory*, Phys. Rev. **D50** (1994) 2700, e-Print: hep-th/9401070.
- [34] V. Frolov and D. Fursaev, *Thermal Fields, Entropy and Black Holes*, Class. Quantum Grav. **15** (1998) 2041.
- [35] T. Jacobson, Phys. Rev. Lett. **75** (1995) 1260, e-Print: hep-th/9504004.
- [36] J. Mäkelä, A. Peltola, Phys. Rev. **D69** (2004) 124008.
- [37] V.E. Hubeny, M. Rangamani, T. Takayanagi, JHEP **0707** (2007) 082, e-Print: arXiv:0705.0016 [hep-th].